

CSCI 5832 Final Project Report (Source Code)

Lucas Derr

May 8, 2024

Anonymous ACL submission

1 Abstract

Fine-tuning a large-language model has become a fundamental skill in natural language processing. These powerful models are malleable and can be fine-tuned to perform specific tasks. This project uses hip-hop lyrics scraped from the internet to fine-tune the distilgpt2 model (a distilled version of GPT-2). The newly trained model is capable of generating text and “auto-completing” user-inputted lyrics. This report will outline all the steps taken to achieve this final result, including data scraping/collection, data cleaning, model training, and evaluation. Finally, as promised, the final model (RhymeBot 1.0) has released an album using some prompts that I have given it. This is included in the appendix.

2 Introduction

Songwriting is a creative process that often requires inspiration and skill. Automating lyric generation could assist musicians who are stuck in a creative block. Additionally, writing lyrics could be considered a ‘human’ task due to its need for significant emotional understanding. It might be difficult or impossible for a mathematical framework such as a transformer to authentically capture the essence of the human experience. Despite this, one cannot ignore the powerful AI models which are still capable of lyric generation can mimic creativity and even novelty. While these types of models will never replace humans, they can be a helpful tool in enhancing the songwriting process.

Building such a model requires a few key steps. First, and most importantly, a dataset must be either generated or downloaded from the internet. To gain experience in web-scraping, a script was written to aggregate a large, customized hip-hop lyric dataset. After the data is gathered and correctly compiled, it is cleaned and prepped so it can be fed into the fine-tuning of the transformer model. After the data

is clean, it is used to retrain the model and adjust its weights. This project elected to use [distilgpt2](#), as it is relatively lightweight and well-regarded in the NLP community. After the model has finished training, it can be evaluated with a combination of objective and subjective metrics. Each of these steps will be explained in greater detail in the Description of Methodology section of the report.

3 Related Work

3.1 GhostWriter: Using an LSTM for Automatic Rap Lyric Generation

<https://aclanthology.org/D15-1221.pdf>

This 2015 paper is very similar to what I will be doing. These researchers used an LSTM to try to generate rap lyrics in the style of a specific artist. Their main metric measures the similarity between the generated lyrics and rhyme density. They conclude that their LSTM model is more effective at generating lyrics than the baseline n-gram model. My project differs because I will use a pre-trained transformer and my dataset will include many artists’ lyrics.

3.2 Hierarchical attention based long short-term memory for Chinese lyric generation

This 2018 paper uses a hierarchical attention model to capture contextual information about the lyrics. The goal with this framework is to capture lyrical style and patterns, maintaining cohesion while generating lyrics. They use cross entropy loss and perplexity to evaluate their model and conclude that the model beats state-of-the-art models, but fails to live up to lyrics written by humans. My project is different because I will not be using a hierarchical attention framework.

075 **3.3 Syllable-level lyrics generation from**
076 **melody exploiting character-level**
077 **language model**

078 This 2024 paper focuses on training a large lan-
079 guage lyric generator model at the syllable level in
080 order to establish a mapping between musical notes
081 and syllables. They elect to use ROUGE, BLEU,
082 and BERT for evaluation but emphasize that there
083 exists no metric that comprehensively captures the
084 quality of generated lyrics. Additionally, they use a
085 rating scale for subjective human evaluation. They
086 conclude that the syllable-based model improves
087 naturality, correctness, and coherence of lyrics’.

088 **4 Lyrics Generation supported by**
089 **Pre-trained Models**

090 This 2022 paper analyzes the generated lyrics of a
091 pre-trained GPT-2 model. They elect to use intrin-
092 sic perplexity to measure the semantic cohesion of
093 the generated text. The paper concludes that the
094 while models generated semantically and syntacti-
095 cally correct lyrics. While this paper emphasizes
096 the analysis of lyrics, my project will be more fo-
097 cused on generating different types of lyrics with
098 different datasets.

099 **5 Methodology**

100 **5.1 Web Scraping**

101 An AI model is only as good as its training data.
102 Due to the complexity and large number of param-
103 eters in transformer models, it is often the case that
104 having more data is always better than having less
105 data. In order to customize the dataset, this project
106 attempted to download the all the lyrics from the
107 entire discography of every artist from IMDB’s
108 50 Best Males Rappers of all Time. The motiva-
109 tion for this is that the ‘best’ rappers should have
110 the highest quality lyrics, which should help the
111 model produce high quality lyrics. The top 50 site
112 is scraped to produce a list of 50 rappers in python.

113 Lyric scraping was done using Python’s Beau-
114 tifulSoup on <https://www.azlyrics.com>. This web-
115 site is ideal because its links to artists’ are easy to
116 access. The process for scraping all the lyrics in-
117 volves two steps: gathering all the links to the lyric
118 pages and gathering the lyrics from the previously
119 generated list of links.

120 Fortunately, the link to every artist page is algo-
121 rithmically generated and straightforward to mimic
122 using just the artist name.

Artist Name	Link to Artist Page
Drake	https://www.azlyrics.com/d/drake.html
2Pac	https://www.azlyrics.com/19/2pac.html

Figure 1: Converting Artist Name to Link

123 Once the links to each artist page has been gener-
124 ated and verified, all the song links within the artist
125 page can be collected and stored in a file. After
126 this process was complete, 11,205 lyric links were
127 generated for 44 different artists.

128 The next step in the process became a consid-
129 erable bottleneck in the progress of this project.
130 Unfortunately, established websites such as az-
131 lyrics.com have bot detection algorithms to prevent
132 people from sending an excessive number of net-
133 work requests. The workaround is to wait around
134 12 seconds between each request. This made
135 the data collection process very time-consuming
136 (would have taken around 48 hours of consecutive
137 runtime). In addition, about halfway through this
138 process, the bot got detected, causing its IP to be
139 banned from future network requests.

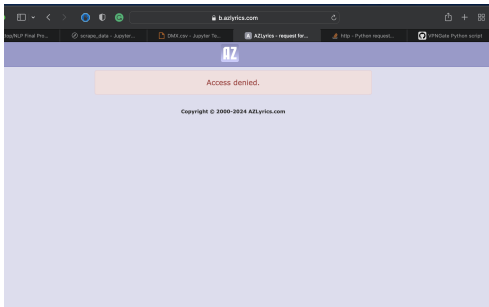


Figure 2: Banned from more Network Requests

140 Eager to continue to ther parts of the project and
141 unwilling to wait for the ban to be lifted, the rest of
142 the artists were scrapped. The final dataset includes
143 22 different artists and 5842 songs, as shown below
144 in 3.

145 **5.2 Data Cleaning**

146 There are several important steps to convert the
147 data from a raw lyric string to a data structure that
148 is suitable for training a transformer model. First,
149 the data is randomly split into two datasets: a train
150 and validation. Train is used to train the model
151 while validation is used to evaluate the model. A
152 small split size of 0.05 was used to maximize the
153 amount of data used to train the model. Afterwards,
154 using the prebuilt tokenizer, the data is converted
155 from text to a list of tokens which are represented

lyrics. Despite challenges such as bot detection algorithms and GPU limitations, the project achieved its primary objective of creating a functional AI model capable of generating and auto-completing hip-hop lyrics.

There are a myriad of ways this model can be improved. Increasing the size of the dataset is an obvious one. Spending \$10 for a month of Colab Pro and increased GPU so that the model can be trained for longer is another. Additionally, due to the GPU limitations and long training times, experimenting with different parameter configurations was not feasible. This could have led to marginal improvements. A point worth candidly mentioning is that not much time was left to spend evaluating the model after it was built. Designing a framework for evaluating the model using different prompts and metrics would have been a great addition to the project. While the ‘songs’ in the appendix pass the eye test, this is hardly a scientific method for evaluating a text generation model, and the evaluation section is the weakest of this report.

In conclusion, the development of RhymeBot 1.0 demonstrates A.I.’s ability to be useful in creative endeavors such as writing lyrics. While AI-generated lyrics may never fully replace the creativity and authenticity of human expression, they can serve as a valuable tool for inspiring and augmenting artistic creation. As technology continues to evolve, the boundaries between human and AI creativity will blur, opening up new possibilities for collaboration and innovation in the world of music and beyond.

7 Appendix

7.1 Song 1

Prompt:

Ayo, y'all other rap bots got no game
I drip hard, don't need the fame
I got the training data to back me up
Your weak rhymes, they're just a hiccup
Your data's shallow, can't stand the heat
While I spit fire, you're just obsolete

Model Output:

Ayo, y'all other rap bots got no game
I drip hard, don't need the fame
I got the training data to back me up
Your weak rhymes, they're just a hiccup
Your data's shallow, can't stand the heat

While I spit fire, you're just obsolete
Your data's no longer safe for the rest of yo
With the same intentions and I'm a dog in this bitch

You're a hypocrite then I'm a sucker
for the bullshit
I'm a motherfucker with the shit
that I put on my head, man
You better watch your mouth,
man, watch your face
You better watch your teeth,
man, watch your mouth
You don't ever see me in my life
Look me in your eyes, man
Look me in your eyes, man
I got plans, for you to get on the set with me
You got plans, for you to get

7.2 Song 2

Prompt:

We out here in Boulder
Reppin hard that rocky mountain kush

Model Output:

We out here in Boulder
Reppin hard that rocky mountain kush

You wanna take me in if you look at me
What you wanna take me
Caught up with this thug gangsta shit

From the East to West Coast
From Cali to L.A.
Rappers know how we ride to a high end

From Cali to L.A.
Rappers know how we ride to a high end

Ay, what you wanna do?
It's all about where you're at
Rappers know where you're at
Rappers know where you're at,
It's all about where you're at
Rappers know where you're at
Rappers know where you're at
I got that gangsta shit I'm bustin on
We gon' bust 'em, we gon' bust 'em

7.3 Song 3

Prompt:

In the city where the streets don't sleep,
 where every corner holds a story deep.
 From broken dreams to shattered hopes,
 this concrete jungle's where I cope.

Model Output:

In the city where the streets don't sleep,
 where every corner holds a story deep.
 From broken dreams to shattered hopes,
 this concrete jungle's where I cope.
 What's the difference?

I know everybody's happy 'fore I die,
 but I know I'm wrong
 I know it's sad how the system do
 I know my mama's gone
 She's always there for the wrong reasons
 Still my baby loves me, but I only know
 I'm not there for me
 Where my baby sees me, we do the same thing
 In the darkness, she feels my pain
 (It's a shame)

Why?
 Why did I feel so damn lost?
 (It's a shame)

What do I do for my people?
 Let me know what we can do for all
 of these people
 We're going all out here, for all
 of these people
 We're going all out here

7.4 Song 4

Prompt:

That's why I fucked yo' bitch,
 you fat motherfucker
 West side, Bad Boy killers get 'em,
 and I'mma shoot em

Model Output:

That's why I fucked yo' bitch,
 you fat motherfucker
 West side, Bad Boy killers get 'em,
 and I'mma shoot em

I'mma shoot the motherfuckin' crip
 And I'mma murder crip, in my hand, hee-ho
 Got him a nine-five, but he like to shoot him

Them dudes like to shoot the motherfucker,
 you fat motherfucker
 Them guns, I got money, so I'mma shoot him
 I'mma shoot 'em, 'cause your boy shoot 'em
 Can't even see him, so what up, let's get away
 I'mma shoot him, I'mma shoot 'em

I wanna see you, but,
 can't believe I can see you
 See I can't see you, but,
 can't believe I can see you
 I got money, so I'mma shoot 'em
 Can't

8 Works Cited

- Potash, Peter, Alexey Romanov, and Anna Rumshisky. "Ghostwriter: Using an lstm for automatic rap lyric generation." Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. 2015.
- Wu, X., Du, Z., Guo, Y. et al. Hierarchical attention based long short-term memory for Chinese lyric generation. Appl Intell 49, 44–52 (2019). <https://doi.org/10.1007/s10489-018-1206-2>
- Zhe Zhang, Karol Lasocki, Yi Yu, and Atsuhiko Takasu. 2024. Syllable-level lyrics generation from melody exploiting character-level language model. In Findings of the Association for Computational Linguistics: EACL 2024, pages 1336–1346, St. Julian's, Malta. Association for Computational Linguistics.
- Rodrigues, M. A., Oliveira, A., Moreira, A., Possi, M. (2022). Lyrics Generation supported by Pre-trained Models. The International FLAIRS Conference Proceedings, 35. <https://doi.org/10.32473/flairs.v35i.13060>